Making crowd estimates is a mix of science, statistics, computer analysis and guesswork

January 31 2017, by Lisa M. Krieger, The Mercury News

It's what everyone wants to know - but is literally impossible to answer: How many people took to the streets Saturday in the women's marches across the United States?

Initial reports in The Associated Press settled for the wildly inconclusive "more than 1 million." But a project this week by two professors to sum up the tallies at 680 marches across the country was pushing estimates as high as 5 million and beyond.

Never before has the size of a crowd been so contentious - yet so darn hard to count.

Crowd counting is an imperfect, laborious and contentious endeavor, ranging from rigorous science to random guesswork. And, as President Donald Trump's weekend fit over the reported size of his inauguration turnout shows, sometimes it's influenced by politics.

Counting swarms of restless humans is like trying to estimate the size of schools of sardine in the sea or herds of wildebeests migrating across the African savanna.

Official Women's March estimates in some cities, such as Oakland, Calif.'s 100,000, relied on aerial photos and analytics. Others enlisted block-by-block head counts. San Jose, Calif., based its 25,000 count on comparisons to similar-sized crowds with reliable counts. San Francisco police, "for reasons of public safety," calculate a tally - but do not

publicly disclose it.

A lot is at stake: Does a cause have widespread popular support, or not?

University of Connecticut professor Jeremy Pressman and University of Denver professor Erica Chenoweth are "crowd sourcing" figures - both high and low estimates - for marches across the world, including 554 in the U.S. As of Monday evening, their tally for U.S. marchers ranged from 3.16 million to 4.68 million. Included are reports of marches where the only source is the marchers themselves - such as 11 marchers in Beaufort, N.C.

"Accurately counting things is hard to do. Producing defensible estimates is hard to do," said Steve Doig, a data journalist expert at Arizona State University, who studies the techniques and challenges of crowd counting. "The only thing you can do is show your work, to show how you did it."

In controlled environments, like a sports stadium, it's easy. "You have turnstiles and tickets sold," said Doig. "But for things that are impromptu or outdoors, with many entry points and many people milling around, further back ... that's much harder."

There are plenty of indirect measures of a crowd: How many people rode subways? How many buses were lined up? How much trash accumulated? How many portable toilets were used?

But direct measurements are far tougher, in sprawling crowds of varying densities. At the Women's March in Washington, D.C., there were brigades of banner-holders, as well as people in wheelchairs and strollers. Crowds spilled out into side streets and alleys. Many children saw the march from the shoulders of their parents. And there were pole and tree-climbers, perched high above the National Mall.

At the same time, Trump lashed out at the media for what he insisted were low estimates at his inauguration.

Three times as many people joined Saturday's march in Washington as Trump's inauguration, according to a new density analysis of seven different YouTube video feeds by Professors Keith Still and Marcel Altenburg of Manchester Metropolitan University in Britain.

The team estimates that at least 470,000 people were on the National Mall for the Women's March. By comparison, about 160,000 people were in the same area in the hour before Trump's swearing-in.

A similar 3-to-1 conclusion was reached in a visual comparison of oblique photographs by Arizona State's Doig.

Still pending are the results of a more detailed - and perhaps more accurate - study by a team led by Curt Westergard, aerial photographer and president of Digital Design & Imaging Service based in Falls Church, Va. Their analysis will be complete by the end of the week.

This team, which has crowd-counted many marches in the mall, used a camera with multiple lenses attached to a tethered balloon, pulled along Pennsylvania Avenue, providing a nearly overhead view of 360-degree images. Their camera captured 2,100 images in four hours. Now, back in the office, they are measuring the density of different parts of the crowd, overlaid on a Google map of topography, and are counting using the same type of automated computer program that counts individual cells in blood specimens.

In some cities, like Los Angeles, estimates can vary by orders of magnitude. Organizers put the number at 750,000. The Los Angeles Police Department said in a statement that "well past" 100,000 people attended but did not provide a more precise number.

Oakland took photos from helicopters, then calculated the total based on crowd density and city maps. In a dense crowd, each person takes 2.5 square feet, said Officer Marco Marquez. Their result: 100,000.

The most significant count is in Washington, D.C. But that's also the most challenging: There are no nearby tall buildings, due to height restrictions; the mall is a "no fly zone," so aircraft or drones can't get images from directly overhead, and the federal government won't provide estimates. And the weather was miserable last weekend, with rain and fog reducing visibility.

While imperfect, all of these approaches to crowd calculations will arrive at a number that is far closer to the truth than that used by participants - or the Trump administration.

"Our goal is not to have a final count," Westergard said. "It is to give a rough order of magnitude. So if somebody says 100,000 and somebody else says 1 million, we can give some sanity to the picture."